

SEQUENCE LISTING

<110> Black Jr., Charles A.

<120> COMPOSITIONS AND METHODS FOR ACTIVATING
GENES OF INTEREST

<130> 5722-2(35722/191928)

<140> 09/446,402

<141> 1999-12-20

<150> PCT/US98/13093

<151> 1998-06-24

<150> 60/050,772

<151> 1997-06-25

<160> 19

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4279

<212> DNA

<213> Artificial Sequence

<220>

<223> Recombinant Molecule containing multiple cloning
site, kozak sequence, LacZ gene.

<221> misc_feature

<222> (1)...(64)

<223> Multiple cloning site

<221> misc_feature

<222> (65)...(79)

<223> Consensus sequence for the "Kozak sequence"
(translation initiation)

<221> prim_transcript

<222> (80)...(4279)

<223> Beta galactosidase

<400> 1

ttaataacgac	tcactatagg	ctagcctcga	gaattcacgc	gtggtacctc	tagagtcgac	60
ccgggcccgc	gccaccatgg	cgcagcacca	tggcctgaaa	taacctctga	aagaggaact	120
tgggttaggta	ccttctgagg	cggaaagaac	cagctgtgga	atgtgtgtca	gttaggggtgt	180
ggaaagtccc	caggctcccc	agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	240
gcaaccagggt	gtggaaagtc	cccaggctcc	ccagcaggca	gaagtatgca	aagcatgcat	300
ctcaattagt	cagcaaccat	agtcgccgcc	ctaactccgc	ccatcccgcc	cctaactccg	360
cccagttccg	cccattctcc	gccccatggc	tgactaattt	tttttattta	tgcagaggcc	420
gaggccgcct	cggcctctga	gctattccag	aagtagtgag	gaggcttttt	tggaggccta	480
ggcttttgca	aaaagcttgg	gatctctata	atctcgcgca	acctattttc	ccctcgaaca	540

ctttttaagc	cgtagataaa	caggctggga	cacttcacat	gagcgaaaaa	tacatcgtca	600
cctgggacat	gttgagatc	catgcacgta	aactcgcaag	ccgactgatg	ccttctgaac	660
aatggaaagg	cattattgcc	gtaagccgtg	gcggtctggt	accggtgggt	gaagaccaga	720
aacagcacct	cgaactgagc	cgcgatatgt	cccagcggtt	caacgcgctg	tatggcgaga	780
tcgatcccgt	cgttttacaa	cgctcgtgact	gggaaaaccc	tggcgttacc	caacttaatc	840
gccttgacgc	acatccccct	ttcgccagct	ggcgtaatag	cgaagaggcc	cgcaaccgatc	900
gcccttccca	acagttgctg	agcctgaatg	gcgaatggcg	ctttgcctgg	tttcgggcac	960
cagaagcggg	gccggaaaag	tggctggagt	gcgatcttcc	tgaggccgat	actgtcgtcg	1020
tcccccaaaa	ctggcagatg	cacggttacg	atgcgcccc	ctacaccaac	gtaacctatc	1080
ccattacggg	caatccgcgc	tttgttccca	cgagagaatcc	gacgggttgt	tactcgtcga	1140
catttaaatgt	tgatgaaagc	tggctacagg	aaggccagac	gcgaattatt	tttgatggcg	1200
ttaaactcggc	gtttcatctg	tgggtgcaacg	ggcgctgggt	cggttacggc	caggacagtc	1260
gtttgcgctc	tgaatttgac	ctgagcgcat	ttttacgcgc	cggagaaaaa	cgcctcgcgg	1320
tgatggtgct	gcgtttggag	gacggcagtt	atctggaaga	tcaggatatg	tggcggatga	1380
gcggcatttt	ccgtgacgtc	tcgttgctgc	ataaacccgac	tacacaaatc	agcgatttcc	1440
atgttgccac	tcgctttaat	gatgatttca	gccgcgctgt	actggaggct	gaagttcaga	1500
tgtgcggcga	gttgcgtagc	tacctacggg	taacagtttc	tttatggcag	ggtgaaacgc	1560
aggctcgccag	cggcaccgcg	cctttcggcg	gtgaaattat	cgatgagcgt	ggtggttatg	1620
cgcgtccggt	cacatacgtg	ctgaacgtcg	aaaacccgaa	actgtggagc	gccgaaatcc	1680
cgaatctcta	tcgtgcggtg	gttgaactgc	acaccgcgga	cggcacgctg	attgaagcag	1740
aagcctgcga	tgtcggtttc	cgcgagggtg	gcattgaaaa	tggctcgtcg	ctgctgaacg	1800
gcaagccgtt	gctgattcga	ggcgtaaac	gtcacgagca	tcctcctctg	catggtcagg	1860
tcattggatga	gcagacgatg	gtgcaggata	tcctgctgat	gaagcagaac	aactttaacg	1920
ccgtgcgctg	ttcgatttat	ccgaaccatc	cgctgtgtgt	cacgctgtgc	gaccgctacg	1980
gcctgtatgt	ggtggatgaa	gccaatattg	aaacccacgg	catggtgcc	atgaatcgtc	2040
tgaccgatga	tcgcgcgtgg	ctaccggcga	tgagcgaacg	cgtaacgcga	atggtgcagc	2100
gcgatcgtaa	tcaccgcaggt	gtgatcatct	ggctgcgtgg	gaatgaatca	ggccacggcg	2160
ctaatacaga	cgcgcgtgat	cgctggatca	aatctgtcga	tccttcccgc	ccggtgcagt	2220
atgaaggcgg	cggagccgac	accacggcca	ccgatattat	ttgcccgatg	tacgcgcgcg	2280
tggatgaaga	ccagcccttc	ccggctgtgc	cgaatggtc	catcaaaaaa	tggctttcgc	2340
tacctggaga	gacgcgccc	ctgacccctt	gcgaatacgc	ccacgcgatg	ggtaacagtc	2400
ttggcgggtt	cgctaaatac	tggcaggcgt	ttcgtcagta	tccccgttta	cagggcgggt	2460
tcgtctggga	ctgggtggat	cagtcgcgtg	ttaaatatga	tgaaaacggc	aaccctggtg	2520
cggcttacgg	cgggtgattt	ggcgatacgc	cgaacgatcg	ccagttctgt	atgaacggtc	2580
tggctctttg	cgaccgcacg	ccgcattccg	cgctgacgga	agcaaaacac	cagcagcagt	2640
ttttccagtt	ccgtttatcc	gggcaaacca	tcgaagtga	cagcgaatac	ctgttccgtc	2700
atagcgataa	cgagctcctg	cactggtatg	tggcgctgga	tggtaagccg	ctggcaagcg	2760
gtgaagtgcc	tctggatgtc	gctccacaag	gtaaacagtt	gattgaactg	cctgaactac	2820
cgcagccgga	gagcgccggg	caactctggc	tcacagtacg	cgtagtgcaa	ccgaacgcga	2880
ccgcattggt	agaagccggg	cacatcagcg	cctggcagca	gtggcgtctg	gcgaaaaacc	2940
tcagtgtgac	gctccccgcc	gcgtcccacg	ccatcccgca	tctgaccacc	agcgaaatgg	3000
atttttgcat	cgagctgggt	aataagcgtt	ggcaatttaa	ccgccagtca	ggctttcttt	3060
cacagatgtg	gattggcgat	aaaaaacaac	tgctgacgcc	gctgcgcgat	cagttcaccc	3120
gtgcaccgct	ggataacgac	attggcgtaa	gtgaagcgac	ccgcattgac	cctaacgcct	3180
gggtcgaacg	ctggaaggcg	gcgggccatt	accaggccga	agcagcgttg	ttgcagtgca	3240
cggcagatac	acttgctgat	gcggtgctga	ttacgaccgc	tcacgcgtgg	cagcatcagg	3300
ggaaaacctt	atttatcagc	cggaaaacct	accggattga	tggtagtggt	caaatggcga	3360
ttaccggttg	tgttgaaagt	gcgagcgata	caccgcattcc	ggcgcggtat	ggcctgaact	3420
gccagctggc	gcaggtagca	gagcgggtaa	actggctcgg	attagggccg	caagaaaact	3480
atcccgcacg	ccttactgcc	gcctgttttg	accgctggga	tctgccattg	tcagacatgt	3540
ataccccgta	cgtcttcccg	agcgaaaacg	gtctgcgctg	cgggacgcgc	gaattgaatt	3600
atggcccaca	ccagtggcgc	ggcgacttcc	agttcaacat	cagccgctac	agtcaacagc	3660
aactgatgga	aaccagccat	cgcatctgct	tgcacgcgga	agaaggcaca	tggctgaata	3720
tcgacggttt	ccatatgggg	attggtggcg	acgactcctg	gagcccgta	gtatcggcgg	3780
aattccagct	gagcgccggg	cgctaccatt	accagttggt	ctggtgtcaa	aaataataat	3840

aaccgggcag	gccatgtctg	cccgatatttc	gcgtaaggaa	atccattatg	tactatttaa	3900
aaaacacaaa	cttttggatg	ttcggtttat	tctttttctt	ttactttttt	atcatgggag	3960
cctacttccc	gtttttcccg	atttggctac	atgacatcaa	ccatatcagc	aaaagtgata	4020
cgggtattat	ttttgccgct	atttctctgt	tctcgctatt	attccaaccg	ctgtttggtc	4080
tgctttctga	caaactcgga	acttgtttat	tgcagottat	aatgggtaca	aataaagcaa	4140
tagcatcaca	aatttcacaa	ataaagcatt	tttttcaactg	cattctagtt	gtggtttgtc	4200
caaactcatc	aatgtatctt	atcatgtctg	gatactctag	agtcgacctg	caggcatgca	4260
agctggcact	ggccgctcgt					4279

<210> 2
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 2	
gaatacaaag cttatgcatg	20

<210> 3
 <211> 13
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 3	
gaatacaaag ctt	13

<210> 4
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 4	
aaagcttatg catgcggccg	20

<210> 5
 <211> 20
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic oligonucleotide

<400> 5	
cggccgcatc tagagggccc	20

<210> 6
 <211> 25

<212> RNA
<213> Unknown

<220>

<223> mRNA sequence for Firefly luciferase

<400> 16

gaauacaaag	cuaaugcaug	cggccgcauc	uagagggccc	ggauccaaau	ggaagacgcc	60
aaaaacauaa	agaaaggccc	ggcgccauuc	uauccucuag	aggauaggac	cgcuaggagag	120
caacugcaua	aggcuauaag	gagauacgcc	cugguuccug	gaacaauugc	uuuuacagau	180
gcacauaucg	aggugaacau	cacguacgcg	gaauacuucg	aaauguccgu	ucgguuggca	240
gaagcuaua	aacgauaugg	gcugaauaca	aaucacagaa	ucgucguaug	cagugaaaac	300
ucucuuaau	ucuuuaugcc	gguguugggc	gccguuuuuu	aucggaguug	caguugcgcc	360
cgcgaagcac	auuuauaaug	aacgugaauu	gcucaacagu	augaacaauu	cgcagccuac	420
cguaguguuu	guuuccaaaa	agggguugca	aaaaauuuug	aacgugcaaa	aaaaauuacc	480
aaauauccag	aaaauuauua	ucauggauuc	uaaaacggau	uaccagggau	uucagucgau	540
guacacguuc	gucacaucuc	aucuaccucc	cgguuuuuuu	gaauacgauu	uuguaccaga	600
guccuuugau	cgugacaaaa	caauugcacu	gauaaugaa	uccucuggau	cuacuggguu	660
accuaagggg	guggcccuuc	cgcauagaac	ugccugcguc	agauucucgc	augccagaga	720
uccuauuuuu	ggcaaucaaa	ucauuccgga	uacugcgauu	uuaaguguug	uuccauucca	780
ucacgguuuu	ggaauguuuu	cuacacucgg	auuuuugauu	uguggauuuc	gagucgucuu	840
aauguaaua	uuugaagaag	agcuguuuuu	acgaucuuu	caggauuaca	aaauucaaa	900
ugcguugcua	guaccaaccc	uauuuucauu	cuucgccaaa	agcacucuga	uugacaaaua	960
cgauuuauuc	aaauuacacg	aaauugcuuc	uggggggcgca	ccucuucgca	aagaagucgg	1020
ggaagcgggu	gcaaaacgcu	uccaucuucc	agggauacga	caaggauaug	ggcucacuga	1080
gacuaacaua	gcuauucuga	uuacacccga	gggggaugau	aaaccggggcg	cggucgguaa	1140
aguuguucca	uuuuuugaag	cgaagguugu	ggaucuggau	accgggaaaa	cgcuggggcg	1200
uaucagaga	ggcgaaauau	gugucagagg	accuauuuu	auguccgguu	auguaaacia	1260
uccggaagcg	accaacgccu	ugaauugaca	ggauugaugg	cuacauucug	gagacauagc	1320
uuacuggggac	gaagacgaac	acuucuucuu	aguugaccgc	uugaagucuu	uaauuaaaua	1380
caaaggauau	cagguggccc	ccgcugaauu	ggaaucgaua	uuguuacaac	acccaacau	1440
cuucgacgcg	ggcguggcag	gucuucccga	cgaugacgcc	ggugaacuuc	ccgccgccgu	1500
uguuguuuug	gagcacggaa	agacgaugac	ggaaaaagag	aucguggauu	acgucgccag	1560
ucaaguaaca	accgcgaaaa	aguugcgcg	aggaguugug	uuuguggacg	aaguaccgaa	1620
aggucuuacc	ggaaaacucg	acgcaagaaa	aaucagagag	auccucauaa	aggccaagaa	1680
gggcgggaa	uccaaaauugu	aaaauguaac	uguauucagc	gaugacgaaa	uucuuagcua	1740
uuguaauccu	ccgagggggc	gagcucccaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1798

<210> 17

<211> 12

<212> DNA

<213> Consensus Kozack sequence

<400> 17

gccgccgcca tg 12

<210> 18

<211> 13

<212> DNA

<213> Consensus Kozack sequence

<400> 18

gccgccrcca ugg 13

<210> 19

<211> 8
<212> RNA
<213> Consensus Shine-Dalgarno sequence

<400> 19
uaaggagg

8

UAGGAGG